

**USE OF E-MAIL FOR CAPTURE OF DOCUMENT METADATA**

**FIELD OF INVENTION**

This invention relates to document capture devices (e.g., document  
5 scanners) that have an inadequate or inconvenient user interface for entering metadata  
values.

**BACKGROUND OF THE INVENTION**

Documents stored into a document repository are most useful when  
10 they have associated metadata. Document metadata consists of information such as  
document title, document author(s), keywords, version, etc. Document repositories  
typically support queries on document metadata.

When a document is captured by a scanning device, the means for  
entering document metadata may be limited. Some scanning devices, or multi-  
15 function devices, lack an adequate User Interface for entering document metadata.  
Some current solutions require that the metadata be specified in advance in a "scan  
template" or a machine readable "paper UI". Therefore, what is needed is a document  
capture device incorporated into a networked system using network tools such as e-  
mail and personal workstations for supplementing the document capture device user  
20 interface.

## SUMMARY OF THE INVENTION

This invention allows the use of standard e-mail systems by a document capture system to solicit document metadata from a user. The document capture system sends an e-mail to the user associated with the captured document. The e-mail “reply to” field specifies the address of the document capture system. The e-mail body contains the names of the metadata properties supported by the repository system for the captured document. An attachment to the e-mail consists of a screen resolution image of the first page of the captured document to remind the user about the particular document for which metadata is being requested. The user interacts with their e-mail access system (e.g., Microsoft Outlook) to generate a reply to the e-mail that contains the original e-mail body. The user then edits the reply e-mail body to supply metadata values for each of the metadata property names. When the user sends the reply, the document capture system extracts the metadata values from the e-mail body and stores the supplied metadata values and the captured document into the repository.

Still other aspects of the present invention will become apparent to those skilled in this art from the following description, wherein there is shown and described an embodiment of this invention by way of illustration of one of the modes best suited to carry out the invention. The invention is capable of other different embodiments and its details are capable of modifications in various, obvious aspects all without departing from the invention. Accordingly, the drawings and descriptions will be regarded as illustrative in nature and not as restrictive.

## BRIEF DESCRIPTION OF THE DRAWINGS

The objects, features and advantages of the invention will become apparent upon consideration of the following detailed disclosure of the invention, especially when it is taken in conjunction with the accompanying drawings wherein:

5           FIGURE 1 is a high level system block diagram illustrating the document capture system in accordance with the present invention;

FIGURE 2 is a high level flowchart showing process steps for the document capture system of FIGURE 1;

10           FIGURE 3 is a graphical illustration showing a first user interface for selecting a template;

FIGURE 4 is a graphical illustration showing a second user interface for response by a user; and

FIGURE 5 is a high level system block diagram illustrating an implementation example using the document capture system of FIGURE 1.

15

## DETAILED DESCRIPTION OF THE INVENTION

FIGURES 1 and 2 depict a high level system block diagram and flowchart illustrating the document capture system in accordance with the present invention. Referring to FIGURES 1 and 2, the document capture system using e-mail for metadata acquisition is accomplished using the following steps. Initially, as shown  
20   in step 101, at a document capture device 20, a user 10 logs in and selects, enters or is associated with basic capture parameters such as a desired repository name, folder name, and metadata source. Once a paper document (not shown) is loaded into a

document feeder of the document capture device 20, the user 10 presses a start button to initiate a scanning operation. When the scanning has completed, the user 10 removes the paper document and returns to their office. Next, the document capture device 20 delivers the basic capture parameters and the scanned document to a Scan Server 22, as shown in step 102. The Scan Server 22 provides temporary document storage while the metadata is acquired from the user 10.

The Scan Server 22 acquires the user's e-mail address and requests a Metadata Acquisition Service 26 to acquire the metadata for this document as shown in step 103. The request includes a URL to the first page of the scanned document, the user's email address, and the target repository's name. The Metadata Acquisition Service 26 requests and receives a capability description from a Repository Manager 24. The capability description includes the names of the supported repositories along with the metadata names supported by each repository. The Metadata Acquisition Service 26 formats an e-mail body with the metadata names supported by the specified repository 32, as shown in step 104.

Next, at step 105, the Metadata Acquisition Service 26 retrieves the first page of the scanned document from the Scan Server 22 and formats it as an email attachment. At step 106, the Metadata Acquisition Service 26 addresses the email to the user 10 and sets the "reply to" fields to the document capture system email account. The formatted email is then sent to an Email Server 28 for delivery to the user 10, as shown in step 106

Turning once again to FIGURES 1 and 2, in step 107 the Email Server 28 notifies the user's email client 30 of the available email and delivers the mail when the user 10 wishes to view it. The user 10 generates a reply in step 108 to the email that includes a copy of the original message. The user 10 edits the message in the reply body by entering metadata values beside the specified metadata names. When the user

10 sends the reply, the Email Server 28 makes it available to the document capture system, as shown in step 109.

Next, in step 110, the Metadata Acquisition Service 26 retrieves the email reply and extracts the metadata values. The Metadata Acquisition Service 26  
5 returns the metadata values to the Scan Server 22 in step 111, completing the request for metadata acquisition. The Scan Server 22 in step 112 makes a request of the Repository Manager 24 to store the document. Next in step 113, the Repository Manager 24 retrieves the captured document from the Scan Server 22. Lastly, the Repository Manager 24 stores the captured document and its associated metadata into  
10 the specified folder in the target repository 32, as shown in step 114.

Referring now to FIGURE 5, a high level system block diagram illustrating an implementation example using the document capture system and flowchart of FIGURES 1 and 2 is shown. A user 10 has a paper document that needs to be shared with a number of colleagues. The user 10 has access to an EOM-branded  
15 document distribution system using the present invention from the Xerox Corporation. To distribute the document, the user 10 logs into a local document capture device, which may be referred to as the Document Centre and selects a personal template that identifies the target repository, folder name, and document format. The user 10 then selects the option to enter document metadata via email, and provides a secure  
20 representation of the user's repository access credentials.

The user 10 interacts with the system using two different user interfaces. The first user interface 36 is at the Document Centre when the user logs in and selects the desired "template". The template contains previously specified document distribution parameters such as repository and folder names, access  
25 credentials, and metadata source. FIGURE 3 shows the user selecting a template. The second user interface 40 is the user's email client. The user 10 replies to an email from the distribution system and supplies metadata values and a distribution list. The

distribution system attached the first page of the scanned document as a reminder of the document to which the metadata applies. FIGURE 4 shows a typical user interface.

The user 10 places the document into the document feeder of the Document Centre, and presses, by way of example only, a green button. The user 10 returns to their office and notices an email from the document distribution system. The email body presents the document metadata names in a form that has space for the user 10 to enter metadata values. There is also a place for the user 10 to specify the document distribution list. The user 10 selects the "Reply" button of the email client and fills in the metadata values and distribution list in the included copy of the original message. When the user 10 selects the "Send" button, the email reply is delivered to the document distribution system, which extracts the metadata values and distribution list. Finally, the document is converted to the specified format, stored into the repository with the specified metadata, and a reference to the document is emailed to the specified distribution list.

More specifically, referring once again to FIGURE 5, the document distribution scenario is accomplished using the following detailed steps. The user 10 logs in and selects a template containing the desired repository name, folder name, and metadata source. Once the paper document is loaded into the Document Centre document feeder, the user 10 presses the green button to initiate the scanning operation. When the scanning has completed, the user 10 removes the paper document and returns to their office. The Document Centre delivers the template and the scanned document in XSM format to a Scan Server 22. The Scan Server 22 notifies a Distribution Service 34 that a new document is ready for distribution. The Distribution Service 34 retrieves a reference to the document and reads the associated template file. From the template, the Distribution Service 34 extracts the repository name and the metadata source that, in this scenario, is the user's email address.

The Distribution Service 34 requests the Metadata Acquisition Service 26 to acquire the metadata for this document. The request includes a URL to the first page of the scanned document, the user's email address, and the target repository's name. The Metadata Acquisition Service 26 requests a capability description from a Repository Orchestrator 24. The capability description includes the names of the supported repositories along with the metadata names supported by each repository. The Metadata Acquisition Service 26 formats an email body with the metadata names supported by the specified repository 32.

The Metadata Acquisition Service 26 retrieves the first page of the scanned document from the Scan Server 22 and formats it as an email attachment. The Metadata Acquisition Service 26 addresses the email to the user 10 and sets the "reply to" field to the distribution system email account. The formatted email is sent to an enterprise Email Server 28 for delivery to the user 10. The Email Server 28 notifies the user's email client of the available email and delivers the mail when the user 10 wishes to view it. The user 10 generates a reply to the email which includes a copy of the original message. The user 10 edits the message in the reply body by entering metadata values for the specified names. Additionally, the user 10 supplies a distribution list to which a reference to the document will be sent. When the user sends the reply, the Email Server 28 makes it available to the distribution system.

The Metadata Acquisition Service 26 retrieves the email reply and extracts the metadata values and distribution list. The Metadata Acquisition Service 26 returns the metadata values and the distribution list to the Distribution Service 34, completing the request for metadata acquisition. The Distribution Service 34 makes a request of the Repository Orchestrator 24 to format and store the document and then email a document reference to a distribution list. The Repository Orchestrator 24 retrieves the scanned document and manages its conversion to the specified format. The converted document and its associated metadata are stored into the specified

folder in the target repository 32. The repository 32 returns a document reference to the Repository Orchestrator 24. The Repository Orchestrator 24 manages the formatting of the document reference into a document token. The document token is attached to an email body and sent to the Email Server 28 for delivery to the distribution list.

While the invention has been described above with reference to specific embodiments thereof, it is apparent that many changes, modifications and variations in the materials, arrangements of parts and steps can be made without departing from the inventive concept disclosed herein. Accordingly, the spirit and broad scope of the appended claims is intended to embrace all such changes, modifications and variations that may occur to one of skill in the art upon a reading of the disclosure. All patent applications, patents and other publications cited herein are incorporated by reference in their entirety.